

Supplement 1 - Neural network model architectures

Model I

Using a 1000 time step input vector of any stimulus frequency, this model predicts:

- response yes/no (0/1)
- the frequency of the stimulus (click, 6, 12, 18, 24, 30 kHz)
- the sound level of the stimulus (5, 10, ..., 95 dB)

Layer (type)	Output Shape	Param #	Connected to
input_3 (InputLayer)	(None, 1000, 1)	0	
batch_normalization_v1_13 (Batch Normalization)	(None, 1000, 1)	4	input_3[0][0]
conv1d_11 (Conv1D)	(None, 1000, 256)	65792	batch_normalization_v1_13[0][0]
batch_normalization_v1_14 (Batch Normalization)	(None, 1000, 256)	1024	conv1d_11[0][0]
activation_11 (Activation)	(None, 1000, 256)	0	batch_normalization_v1_14[0][0]
conv1d_12 (Conv1D)	(None, 1000, 128)	4194432	activation_11[0][0]
batch_normalization_v1_15 (Batch Normalization)	(None, 1000, 128)	512	conv1d_12[0][0]
activation_12 (Activation)	(None, 1000, 128)	0	batch_normalization_v1_15[0][0]
average_pooling1d (Average Pooling1D)	(None, 250, 128)	0	activation_12[0][0]
dropout_5 (Dropout)	(None, 250, 128)	0	average_pooling1d[0][0]
conv1d_13 (Conv1D)	(None, 250, 64)	524352	dropout_5[0][0]
batch_normalization_v1_16 (Batch Normalization)	(None, 250, 64)	256	conv1d_13[0][0]
activation_13 (Activation)	(None, 250, 64)	0	batch_normalization_v1_16[0][0]
conv1d_14 (Conv1D)	(None, 250, 32)	65568	activation_13[0][0]
batch_normalization_v1_17 (Batch Normalization)	(None, 250, 32)	128	conv1d_14[0][0]
activation_14 (Activation)	(None, 250, 32)	0	batch_normalization_v1_17[0][0]
average_pooling1d_1 (Average Pooling1D)	(None, 125, 32)	0	activation_14[0][0]
dropout_6 (Dropout)	(None, 125, 32)	0	average_pooling1d_1[0][0]
conv1d_15 (Conv1D)	(None, 125, 16)	8208	dropout_6[0][0]
batch_normalization_v1_18 (Batch Normalization)	(None, 125, 16)	64	conv1d_15[0][0]
activation_15 (Activation)	(None, 125, 16)	0	batch_normalization_v1_18[0][0]
conv1d_16 (Conv1D)	(None, 125, 8)	1032	activation_15[0][0]
batch_normalization_v1_19 (Batch Normalization)	(None, 125, 8)	32	conv1d_16[0][0]
activation_16 (Activation)	(None, 125, 8)	0	batch_normalization_v1_19[0][0]
average_pooling1d_2 (Average Pooling1D)	(None, 62, 8)	0	activation_16[0][0]
dropout_7 (Dropout)	(None, 62, 8)	0	average_pooling1d_2[0][0]
conv1d_17 (Conv1D)	(None, 62, 4)	132	dropout_7[0][0]
batch_normalization_v1_20 (Batch Normalization)	(None, 62, 4)	16	conv1d_17[0][0]
activation_17 (Activation)	(None, 62, 4)	0	batch_normalization_v1_20[0][0]
conv1d_18 (Conv1D)	(None, 62, 1)	9	activation_17[0][0]
batch_normalization_v1_21 (Batch Normalization)	(None, 62, 1)	4	conv1d_18[0][0]
activation_18 (Activation)	(None, 62, 1)	0	batch_normalization_v1_21[0][0]
flatten_2 (Flatten)	(None, 62)	0	activation_18[0][0]

dense_2 (Dense)	(None, 32)	2016	flatten_2[0][0]
main_prediction (Dense)	(None, 1)	33	dense_2[0][0]
frequency_prediction (Dense)	(None, 6)	198	dense_2[0][0]
sl_prediction (Dense)	(None, 20)	660	dense_2[0][0]
=====			
Total params: 4,864,472			
Trainable params: 4,863,452			
Non-trainable params: 1,020			

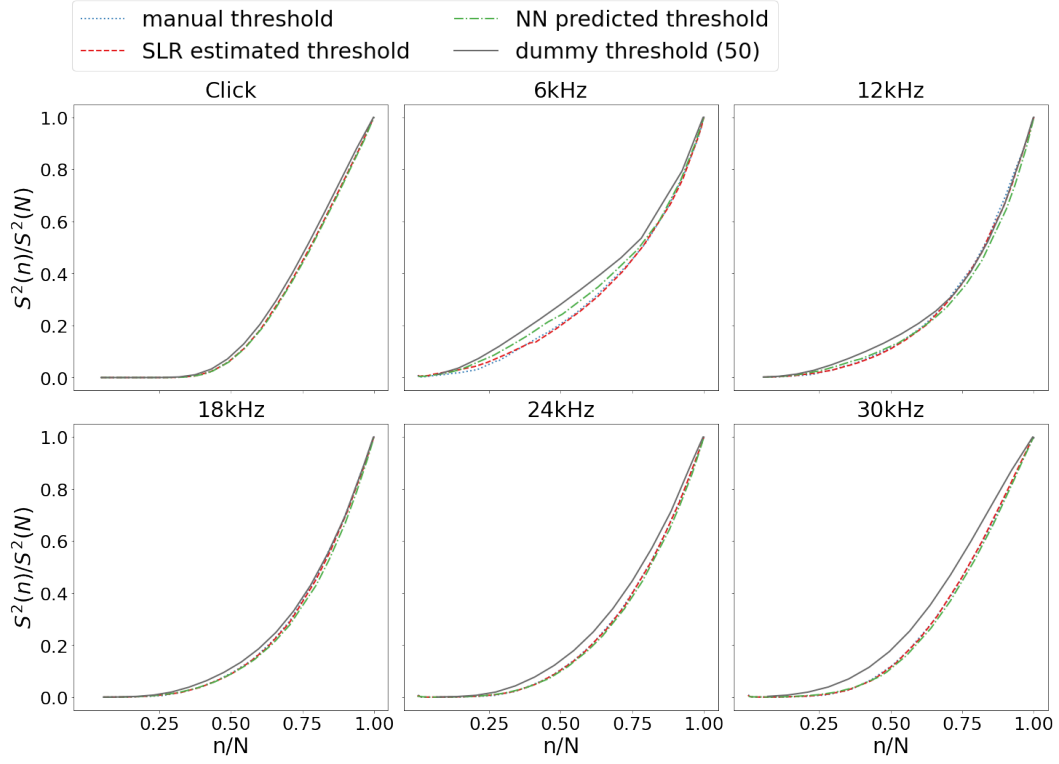
Model II

Using an input vector of 20 sound level prediction scores from model I output, this model predicts:

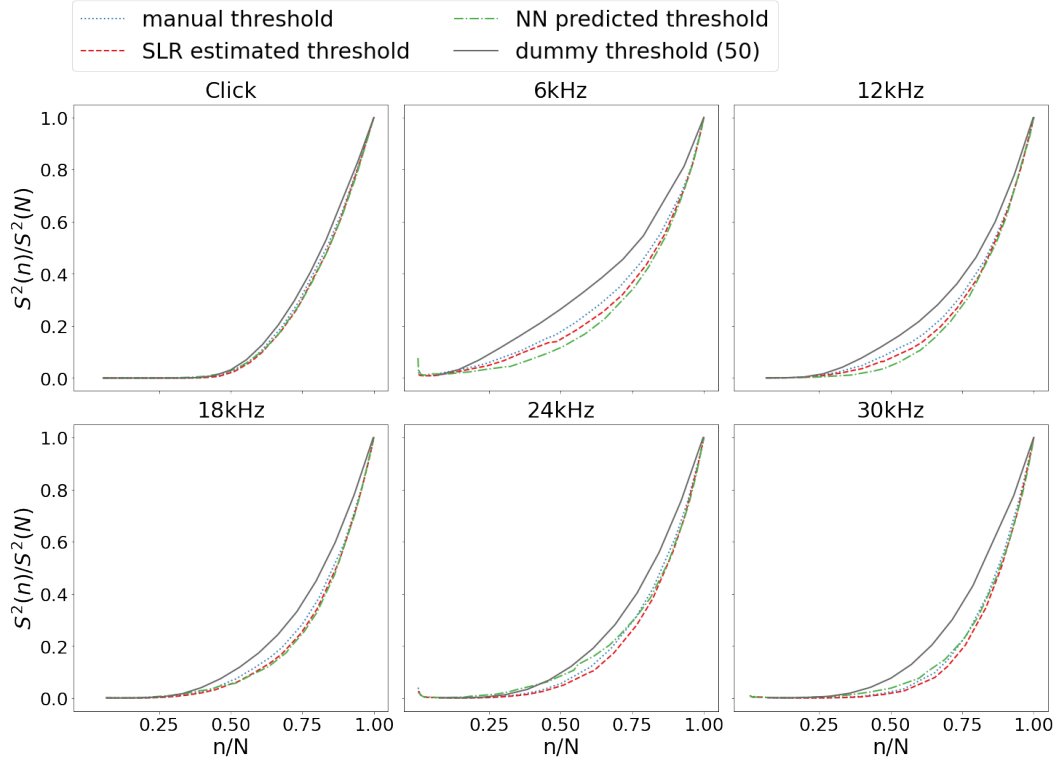
- the frequency of the stimulus (click, 6, 12, 18, 24, 30 kHz)
- the hearing threshold (5, 10, ..., 95 dB)

Layer (type)	Output Shape	Param #	Connected to
=====			
input_2 (InputLayer)	(None, 20, 1)	0	
batch_normalization_v1_4 (Batch Normalization)	(None, 20, 1)	4	input_2[0][0]
conv1d_3 (Conv1D)	(None, 20, 128)	896	batch_normalization_v1_4[0][0]
batch_normalization_v1_5 (Batch Normalization)	(None, 20, 128)	512	conv1d_3[0][0]
activation_3 (Activation)	(None, 20, 128)	0	batch_normalization_v1_5[0][0]
conv1d_4 (Conv1D)	(None, 20, 64)	41024	activation_3[0][0]
batch_normalization_v1_6 (Batch Normalization)	(None, 20, 64)	256	conv1d_4[0][0]
activation_4 (Activation)	(None, 20, 64)	0	batch_normalization_v1_6[0][0]
max_pooling1d_1 (MaxPooling1D)	(None, 6, 64)	0	activation_4[0][0]
dropout_2 (Dropout)	(None, 6, 64)	0	max_pooling1d_1[0][0]
conv1d_5 (Conv1D)	(None, 6, 32)	8224	dropout_2[0][0]
batch_normalization_v1_7 (Batch Normalization)	(None, 6, 32)	128	conv1d_5[0][0]
activation_5 (Activation)	(None, 6, 32)	0	batch_normalization_v1_7[0][0]
conv1d_6 (Conv1D)	(None, 6, 16)	1552	activation_5[0][0]
batch_normalization_v1_8 (Batch Normalization)	(None, 6, 16)	64	conv1d_6[0][0]
activation_6 (Activation)	(None, 6, 16)	0	batch_normalization_v1_8[0][0]
flatten_1 (Flatten)	(None, 96)	0	activation_6[0][0]
dense_1 (Dense)	(None, 64)	6208	flatten_1[0][0]
main_prediction (Dense)	(None, 21)	1365	dense_1[0][0]
frequency_prediction (Dense)	(None, 6)	390	dense_1[0][0]
=====			
Total params: 60,623			
Trainable params: 60,141			
Non-trainable params: 482			

Supplement 2 - Evaluation curves



(a) NN/SLR GMC-ING (experiments 2 and 6)



(b) NN/SLR ING-GMC (experiments 3 and 7)

Fig. S1: Objective comparison of the quality of threshold finding methods using evaluation curves. Evaluation curves allow the relative comparison of threshold finding methods without requiring absolute ground truth labels. Here, four methods are compared: manual thresholds (blue, dotted lines), SLR estimations (red, dashed lines), NN predictions (green, dash-dotted lines), and a “always 50 dB” dummy method (grey, solid lines).

Fig. S1: (continued from previous page)

Separate plots show evaluation curves for each stimulus (click, 6 kHz - 30 kHz). Plots show the normalized time variance of the averaged signal $S^2(n)/S^2(N)$ (y-axis) vs. the total percentage of ABR curves included in the cumulative average n/N (x-axis). a) shows NN predictions and SLR estimations from experiments 2 and 6, b) shows NN predictions and SLR estimations from experiments 3 and 7, as introduced in table 2. Two methods can be compared in a way that the evaluation curve of the better method deviates from zero later. Ideally, the curve of the best method is always below all other curves. Strangely, when GMC-trained/calibrated models were tested on ING data (experiments 2 and 6), human, NN and SLR were not much better than the dummy method. In addition, except for the 6 kHz stimulus, they did not differ much from each other. In contrast, when ING-trained/calibrated models were tested on GMC data (experiments 3 and 7), NN and SLR models were mostly better than the human method, with NN being best for click and 6 kHz to 18 kHz and SLR being best for 24 kHz and 30 kHz.

Supplement 3 - Volcano plots of GMC mutant lines

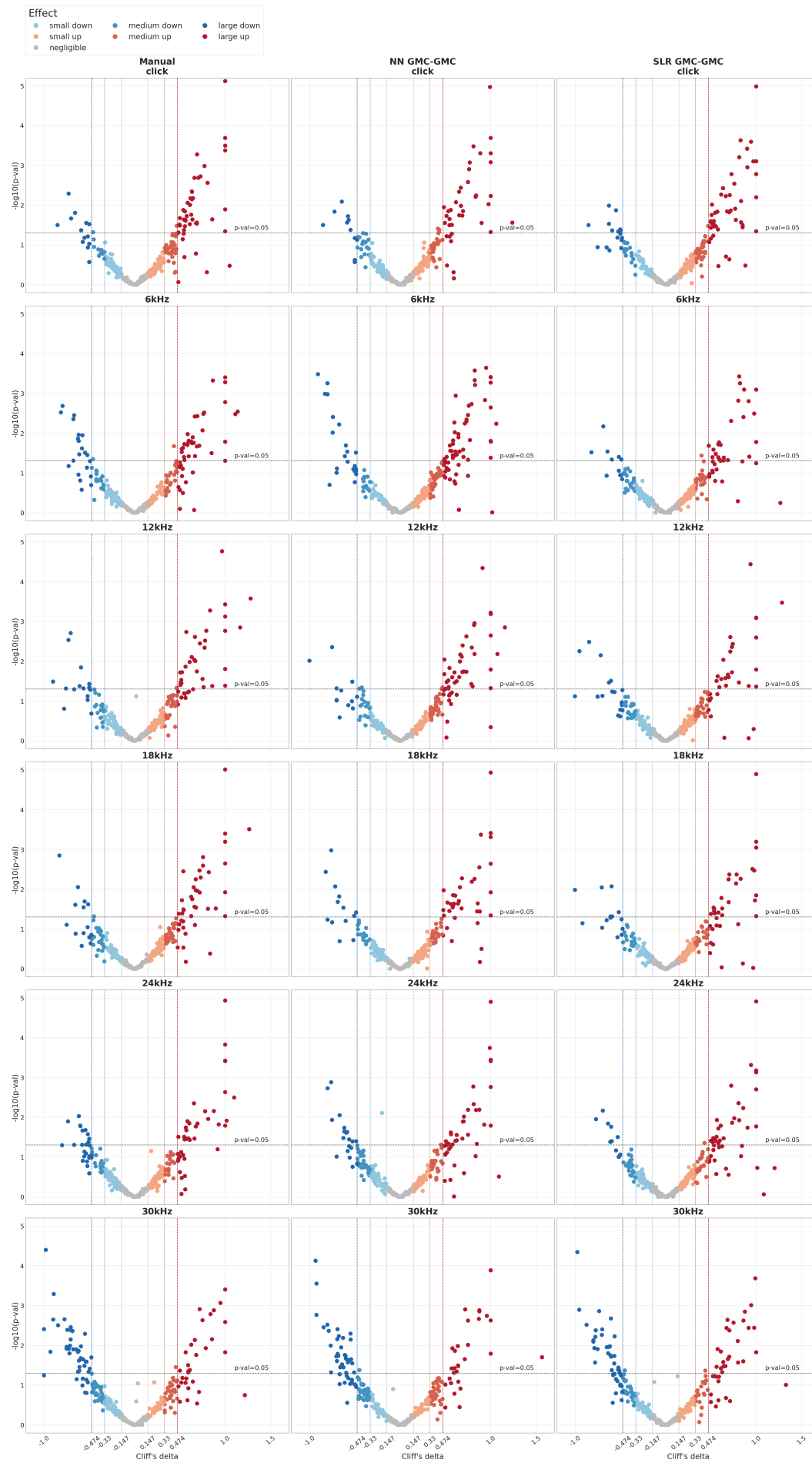


Fig. S2: (continued from previous page)

Biologically relevant changes in hearing thresholds - GMC lines, all stimuli. For each GMC line, Volcano plots show significance vs. relevance for comparisons of mutant and control mice. For each mouse line, represented by a dot, hearing thresholds were used to calculate significance (Wilcoxon test, y-axis) and non-parametric effect size (Cliff's delta [39], x-axis) of mutant vs. control animals. Vertical lines indicate margins for small (0.147), medium (0.33) and large (0.474) effects as suggested in [38]. The horizontal line indicates the 0.05 significance threshold level. Accordingly, mutant lines represented by data points in the upper left and upper right areas denote lines with significant as well as relevant changes and thus are considered worthwhile candidates (see supplement table S1). Dot colors in addition represent effect size as shown in the legend. Plot rows represent different stimuli (click, 6 kHz - 30 kHz). Columns compare the three hearing threshold finding methods compared in this work (left: manual, middle: NN, right: SLR).

Supplement 4 - Comparison of top click candidate genes with hearing threshold changes

Table S1: Comparison of top candidate genes with modified click hearing threshold. The columns show genes with a strong effect size for mutant-control group comparisons of click hearing thresholds. In each column, genes are ordered by non-parametric effect size (Cliff's delta [39]), in descending order. Only genes with a significant ($p < 0.05$) and large effect size ($|d| > 0.474$, [38]) in any of the three methods are shown. Column headers indicate the direction of the effect and the method that was used to identify the hearing thresholds, respectively. The “combined” column contains the combined list of all three methods.

increased click threshold				decreased click threshold			
combined	manual	NN	SLR	combined	manual	NN	SLR
<i>Palm3</i>	<i>Zfp280d</i>	<i>Slc20a2</i>	<i>Lsm1</i>	<i>Dio1</i>	<i>Dio1</i>	<i>Dio1</i>	<i>Dio1</i>
<i>Zfp280d</i>	<i>Lsm1</i>	<i>Prkd2</i>	<i>Strbp</i>	<i>Gstm6</i>	<i>Gstm6</i>	<i>Ucp1</i>	<i>Hnf4a</i>
<i>Strbp</i>	<i>Strbp</i>	<i>Zfp280d</i>	<i>Palm3</i>	<i>Hepacam2</i>	<i>Hepacam2</i>	<i>Cilp2</i>	<i>Gstt1</i>
<i>Hipk3</i>	<i>Hipk3</i>	<i>Strbp</i>	<i>Zfp280d</i>	<i>Slc25a15</i>	<i>Slc25a15</i>	<i>Gstm6</i>	<i>Ostf1</i>
<i>Prkd2</i>	<i>Palm3</i>	<i>Lsm1</i>	<i>Hipk3</i>	<i>Ostf1</i>	<i>Ostf1</i>	<i>Raet1c</i>	<i>Angptl3</i>
<i>Lsm1</i>	<i>Prkd2</i>	<i>Palm3</i>	<i>Prkd2</i>	<i>Ucp1</i>	<i>Rab35</i>	<i>Rab35</i>	<i>Phactr4</i>
<i>Mipol1</i>	<i>Mipol1</i>	<i>Hipk3</i>	<i>Plag1</i>	<i>Rab35</i>	<i>Cilp2</i>	<i>Slc25a15</i>	<i>Rab35</i>
<i>Hoxa2</i>	<i>Hoxa2</i>	<i>Mipol1</i>	<i>Hoxa2</i>	<i>Hnf4a</i>			<i>Cilp2</i>
<i>Slc20a2</i>	<i>Plag1</i>	<i>Vps13c</i>	<i>Aspa</i>	<i>Cilp2</i>			<i>Prox2</i>
<i>Plag1</i>	<i>Ldlr</i>	<i>Plag1</i>	<i>Dpp3</i>	<i>Prox2</i>			
<i>Ldlr</i>	<i>Hunk</i>	<i>Nacc1</i>	<i>Vps13c</i>	<i>Phactr4</i>			
<i>Hunk</i>	<i>Atp5g2</i>	<i>Btbd9</i>	<i>Mipol1</i>	<i>Raet1c</i>			
<i>Atp5g2</i>	<i>Ube3c</i>	<i>Alkbh6</i>	<i>Btbd9</i>	<i>Gstt1</i>			
<i>Btbd9</i>	<i>Btbd9</i>	<i>Zdhhc5</i>	<i>Alkbh6</i>	<i>Angptl3</i>			
<i>Ube3c</i>	<i>Nacc1</i>	<i>Rabgap1</i>	<i>Zdhhc5</i>				
<i>Nacc1</i>	<i>Tle1</i>	<i>Aspa</i>	<i>Bms1</i>				
<i>Tle1</i>	<i>Gpatch2l</i>	<i>Hoxa2</i>	<i>Nacc1</i>				
<i>Gpatch2l</i>	<i>Cidec</i>	<i>Sytl4</i>	<i>Plekha1</i>				
<i>Cidec</i>	<i>Alkbh6</i>	<i>Pdcd5</i>	<i>Ldlr</i>				
<i>Alkbh6</i>	<i>Spryd3</i>	<i>Ldlr</i>	<i>Adprm</i>				
<i>Spryd3</i>	<i>Bccip</i>	<i>Tll12</i>	<i>Wrnip1</i>				
<i>Bccip</i>	<i>Pdcd5</i>	<i>Hunk</i>	<i>Sytl4</i>				
<i>Pdcd5</i>	<i>Ppp4r3b</i>	<i>Ppp4r3b</i>	<i>Hdac1</i>				
<i>Ppp4r3b</i>	<i>Csnk1g2</i>	<i>Kansl1l</i>	<i>Ppp4r3b</i>				
<i>Csnk1g2</i>	<i>Ngdn</i>	<i>Bccip</i>	<i>Tle1</i>				
<i>Ngdn</i>	<i>Gpsm2</i>	<i>Csnk1g2</i>	<i>Gpsm2</i>				
<i>Gpsm2</i>	<i>Aspa</i>	<i>Plekha1</i>	<i>Pkn2</i>				
<i>Aspa</i>	<i>Plekha1</i>	<i>Tle1</i>	<i>Me2</i>				
<i>Pfkfb3</i>	<i>Pfkfb3</i>	<i>Fdx1</i>	<i>Tll12</i>				
<i>Plekha1</i>	<i>Wrnip1</i>	<i>Gpsm2</i>	<i>Rabgap1</i>				
<i>Wrnip1</i>	<i>Zdhhc5</i>	<i>Atp5g2</i>	<i>Csnk1g2</i>				
<i>Zdhhc5</i>	<i>Uggt2</i>		<i>Tbl1xr1</i>				
<i>Dpp3</i>	<i>Gdi2</i>		<i>Pdcd5</i>				
<i>Uggt2</i>	<i>Sec14l4</i>		<i>Hunk</i>				
<i>Gdi2</i>	<i>Tanc2</i>		<i>Atp5g2</i>				
<i>Sec14l4</i>	<i>Rfxank</i>		<i>Fdx1</i>				
<i>Sytl4</i>	<i>Me2</i>		<i>Tanc2</i>				
<i>Rfxank</i>	<i>Gsk3a</i>		<i>Cenpv</i>				
<i>Tanc2</i>			<i>Bccip</i>				
<i>Me2</i>							
<i>Gsk3a</i>							
<i>Rabgap1</i>							
<i>Tll12</i>							
<i>Fdx1</i>							
<i>Bms1</i>							
<i>Hdac1</i>							
<i>Tbl1xr1</i>							
<i>Adprm</i>							
<i>Vps13c</i>							
<i>Kansl1l</i>							
<i>Pkn2</i>							
<i>Cenpv</i>							

Supplement 5 - Comparison of top 30 kHz candidate genes with hearing threshold changes

Table S2: Comparison of top candidate genes with modified 30 kHz hearing threshold. The columns show genes with a strong effect size for mutant-control group comparisons of 30 kHz hearing thresholds. In each column, genes are ordered by non-parametric effect size (Cliff's delta [39]), in descending order. Only genes with a significant ($p < 0.05$) and large effect size ($|d| > 0.474$, [38]) in any of the three methods are shown. Column headers indicate the direction of the effect and the method that was used to identify the hearing thresholds, respectively. The “combined” column contains the ranked combined list of all three methods.

increased 30 kHz threshold				decreased 30 kHz threshold			
combined	manual	NN	SLR	combined	manual	NN	SLR
<i>Chst5</i>	<i>Zfp280d</i>	<i>Chst5</i>	<i>Zfp280d</i>	<i>Miga1</i>	<i>Miga1</i>	<i>Gfpt2</i>	<i>Gfpt2</i>
<i>Palm3</i>	<i>Strbp</i>	<i>Zfp280d</i>	<i>Plag1</i>	<i>Gfpt2</i>	<i>Gfpt2</i>	<i>Rest</i>	<i>Hnf4a</i>
<i>Strbp</i>	<i>Palm3</i>	<i>Strbp</i>	<i>Strbp</i>	<i>Aqp6</i>	<i>Aqp6</i>	<i>Tap1</i>	<i>Rest</i>
<i>Zfp280d</i>	<i>Uggt2</i>	<i>Plag1</i>	<i>Ttll12</i>	<i>Rest</i>	<i>Rest</i>	<i>Hnf4a</i>	<i>Fgfr1op</i>
<i>Uggt2</i>	<i>Prkd2</i>	<i>Ube3c</i>	<i>Ube3c</i>	<i>Cotl1</i>	<i>Cotl1</i>	<i>Becn1</i>	<i>Msh5</i>
<i>Prkd2</i>	<i>Ube3c</i>	<i>Prkd2</i>	<i>Prkd2</i>	<i>Hnf4a</i>	<i>Hnf4a</i>	<i>Fbp2</i>	<i>Dis3</i>
<i>Ube3c</i>	<i>Plag1</i>	<i>Palm3</i>	<i>Palm3</i>	<i>Cenph</i>	<i>Cenph</i>	<i>Msh5</i>	<i>Miga1</i>
<i>Plag1</i>	<i>Nacc1</i>	<i>Uggt2</i>	<i>Mipol1</i>	<i>Dnajc27</i>	<i>Msh5</i>	<i>Aqp6</i>	<i>Fbp2</i>
<i>Nacc1</i>	<i>Gsk3a</i>	<i>Ttll12</i>	<i>Mthfsl</i>	<i>Msh5</i>	<i>Dnajc27</i>	<i>Etfldh</i>	<i>Becn1</i>
<i>Gsk3a</i>	<i>Hipk3</i>	<i>Gsk3a</i>	<i>Cnot6l</i>	<i>Lama1</i>	<i>Lama1</i>	<i>Ostf1</i>	<i>Dnajc27</i>
<i>Hipk3</i>	<i>Ttll12</i>	<i>Hipk3</i>	<i>Gsk3a</i>	<i>Pkig</i>	<i>Pkig</i>	<i>Cilp2</i>	<i>Ppy</i>
<i>Ttll12</i>	<i>Zdhhc5</i>	<i>Nacc1</i>	<i>Csnk1g2</i>	<i>Fgfr1op</i>	<i>Fgfr1op</i>	<i>Lss</i>	<i>Tap1</i>
<i>Zdhhc5</i>	<i>Ldlr</i>	<i>Rnf186</i>	<i>Rnf186</i>	<i>Fbp2</i>	<i>Fbp2</i>	<i>Gstm6</i>	<i>Fam162a</i>
<i>Mipol1</i>	<i>Ngdn</i>	<i>Adprm</i>	<i>Adprm</i>	<i>Cyp7a1</i>	<i>Cyp7a1</i>	<i>Pkig</i>	<i>Cenph</i>
<i>Ldlr</i>	<i>Adprm</i>	<i>Zdhhc5</i>	<i>Hipk3</i>	<i>Fam162a</i>	<i>Fam162a</i>	<i>Galk2</i>	<i>Cyp7a1</i>
<i>Ngdn</i>	<i>Pdcd5</i>	<i>Aldh1l1</i>	<i>Alkbh6</i>	<i>Galk2</i>	<i>Galk2</i>	<i>Mgat1</i>	<i>Galk2</i>
<i>Adprm</i>	<i>Rnf186</i>	<i>Csnk1g2</i>	<i>Aldh1l1</i>	<i>Yae1d1</i>	<i>Yae1d1</i>	<i>Slc25a15</i>	<i>Lss</i>
<i>Rnf186</i>	<i>Fabp2</i>	<i>Paox</i>	<i>Pdcd5</i>	<i>Becn1</i>	<i>Becn1</i>	<i>Ppy</i>	<i>Cotl1</i>
<i>Pdcd5</i>	<i>Entpd1</i>	<i>Cnot6l</i>	<i>Uggt2</i>	<i>Ly6g6d</i>	<i>Ly6g6d</i>	<i>Ctc1</i>	<i>Pdhx</i>
<i>Fabp2</i>	<i>Csnk1g2</i>	<i>Alkbh6</i>	<i>Nacc1</i>	<i>Jmjd6</i>	<i>Jmjd6</i>	<i>Sarnp</i>	<i>Vwa8</i>
<i>Aldh1l1</i>		<i>Fabp2</i>	<i>Paox</i>	<i>Vwa8</i>	<i>Dis3</i>	<i>Vwa8</i>	<i>Cilp2</i>
<i>Entpd1</i>			<i>Zdhhc5</i>	<i>Dis3</i>	<i>Vwa8</i>	<i>Cyp7a1</i>	<i>Zbtb24</i>
<i>Cnot6l</i>			<i>Gpsm2</i>	<i>Tap1</i>	<i>Tap1</i>	<i>Dnajc27</i>	<i>Mgat1</i>
<i>Csnk1g2</i>			<i>Npepps</i>	<i>Phactr4</i>	<i>Phactr4</i>	<i>Fgfr1op</i>	<i>Yae1d1</i>
<i>Paox</i>				<i>C1galt1</i>	<i>C1galt1</i>	<i>Nubp2</i>	<i>Lactb</i>
<i>Mthfsl</i>				<i>Cth</i>	<i>Cth</i>	<i>Cotl1</i>	<i>Sarnp</i>
<i>Gpsm2</i>				<i>Cilp2</i>	<i>Cilp2</i>	<i>Fam162a</i>	<i>Scmh1</i>
<i>Npepps</i>				<i>Ppy</i>	<i>Ppy</i>	<i>Pdhx</i>	<i>Rcn1</i>
<i>Alkbh6</i>				<i>Cdc123</i>	<i>Cdc123</i>	<i>Lactb</i>	<i>Nubp2</i>
				<i>Scmh1</i>	<i>Scmh1</i>	<i>Capn12</i>	
				<i>Ndufa10</i>	<i>Ndufa10</i>	<i>Yae1d1</i>	
				<i>Pdhx</i>	<i>Pdhx</i>	<i>Jmjd6</i>	
				<i>Lss</i>	<i>Lss</i>		
				<i>Capn12</i>	<i>Capn12</i>		
				<i>Sarnp</i>	<i>Sarnp</i>		
				<i>Cystm1</i>	<i>Cystm1</i>		
				<i>Ube3a</i>	<i>Ube3a</i>		
				<i>Nubp2</i>			
				<i>Slc25a15</i>			
				<i>Zbtb24</i>			
				<i>Rcn1</i>			
				<i>Etfldh</i>			
				<i>Ostf1</i>			
				<i>Lactb</i>			
				<i>Ctc1</i>			
				<i>Mgat1</i>			
				<i>Gstm6</i>			